

Putative pathogenicity genes of *Phytophthora cinnamomi* identified via RNA-Seq analysis of pre-infection structures

Anandi Reitmann, Dave K. Berger, and Noelani van den Berg

A. Reitmann

University of Pretoria, Department of Genetics, Forestry and Agricultural Biotechnology Institute, Pretoria, 0002, - outh Africa

D.K. Berger

University of Pretoria, Department of Plant - cience, Forestry and Agricultural Biotechnology Institute, Pretoria, 0002, - outh Africa

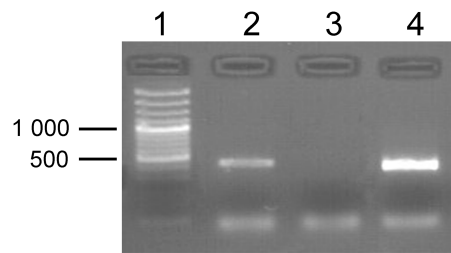
N. van den Berg (Corresponding author)

University of Pretoria, Department of Microbiology and Plant Pathology, Forestry and Agricultural Biotechnology Institute, Pretoria, 0002, - outh Africa

E-mail: noelani.vdberg@fabi.up.ac.za

Tel. +27 12 420 3856

Online Resources



Online Resource 1. Species-specific PCR confirming the identity of isolate CNVDB29 as *Phytophthora cinnamomi*. A primer pair targeting the *LPV3* gene (Kong et al. 2003) led to the successful amplification of a PCR product of 450 bp in size, confirming the identity of isolate CNVDB29 as *P. cinnamomi*. 1) O'GeneRuler 100 bp DNA ladder 2) *Phytophthora cinnamomi* (CMW33388) as a positive control 3) Negative control 4) CNVDB29.

Online Resource 2. Transcripts from the *Phytophthora cinnamomi* CNV- B29 cyst and germinating cyst c- NA library selected for expression analysis. The best *Phytophthora* protein match during blastx analysis (E-value $<1 \times 10^{-5}$) and its functional description is indicated for each transcript.

Unigene ID	Protein ID of best blast hit	Species	Putative function	Bit Score	E-value
C83210	Phyra_97565 ^a	<i>P. ramorum</i>	Pr_Avh387	32.73	1.42E-06
U15350	PITG_19936	<i>P. infestans</i>	NPP1-like protein	111.69	5.04E-26
U4987	PITG_06370	<i>P. infestans</i>	elicitin INL3a-like protein	55.84	3.23E-09
U6179	PITG_03640	<i>P. infestans</i>	CBEL putative	167.16	5.69E-47
U6659	AAN05466.1 ^b	<i>P. cinnamomi</i>	polygalacturonase	192.2	1.14E-57
U6809	PITG_16183	<i>P. infestans</i>	adhesin-like protein	200.29	3.25E-52
U8949	PITG_09760	<i>P. infestans</i>	mucin-like protein	157.15	8.95E-39

SNEL2 - suppressor of necrosis 1-like protein 2, avh – avirulence homologue, NPP1 – necrosis-inducing *Phytophthora* protein 1, CBEL – cellulose binding elicitor lectin

^aProtein ID numbers from FungiDB database version 2.3 (March, 2014)

^bProtein ID numbers from NCBI database

PITG ID numbers are from the *Phytophthora infestans* Sequencing Project, Broad Institute of Harvard and MIT (<http://www.broadinstitute.org/>)

Online Resource 3. Sequences of primers used in this study, targeting putative pathogenicity related genes of *Phytophthora cinnamomi* CNVDB29.

Target	Forward Primer 5'-3'	Reverse primer 5'-3'	PCR product (bp)	Ta (°C)
C83210	TACCGAGGCTGAAAATCC	CTTCTTGTCGTCGTTAC	88	58
U15350	GTTCCAAGACCTCATCAC	TCTGTATCCTTCTCTCC	102	58
U4987	ACAACGGGCATATTCAAG	GTTAGTAGCTCCCATCAG	75	58
U6179	TGGTACTCGTTGAAGAAAG	AAGACTGTGTACGGTTTG	92	58
U6659	GTCACCATCACAGGAAAC	TGCTGAAGACCGTGTTAG	79	58
U6809	TCATCCAAGAAGTCATCG	TTCCAGTTCAGTCAGTTC	78	58
U8949	GCTCCAGTCTTGATCTTC	CCACGAGAACCTGTCTAG	84	58
Primers directed to candidate reference genes from the literature:				
Ubc ^a	CTGAACATCTACTCCCGGCC	CGTTGGCATTGATGTTGCAG	70-150	61
β -tubulin ^a	TGACCCAGCAGCAGTTCG	CATCGACCTCCTTCGTGCTC	117	61
WS041 ^b	GCTGACCAACAACGGCACG	GATCACCTCCGTACCTGCG	178	61

^a(King et al. 2010)

^b(Narayan et al. 2010)

Alignment of jgi-Phyci1-30885 ORF from the *Phytophthora cinnamomi* genome with the sequenced nucleotide data

```

jgi-Phyci1-30885      1  atgcgcggtgttgctgtgggtgctgctcgtggccttgggtcaccttgctatc
Sequenced 30885      1  -----

jgi-Phyci1-30885     51  ctcgactgatgctttgagcacgaacaactcggacaagaaacaggtggtac
Sequenced 30885     41  .....r.

jgi-Phyci1-30885    101  agcctaacagcgaggaggtggcgacacgtatgctcgtgctaactacgag
Sequenced 30885     91  .....

jgi-Phyci1-30885    151  aacaacaacgacaagcggatgctgagaggtgaaagcaagatgacctacgc
Sequenced 30885    141  .....

jgi-Phyci1-30885    201  cacgaacgcgaaaacaacgaagagcgcggtcttcatcaaattccaag
Sequenced 30885    191  .....

jgi-Phyci1-30885    251  gttccgtccagagactgagggagaaaactccgtgacattaaccggttttcg
Sequenced 30885    241  .....

jgi-Phyci1-30885    301  acgaaaaacattgaaaagcgatttcaaaagttggccgacaaaaggagaac
Sequenced 30885    291  .....g.....

jgi-Phyci1-30885    351  ccccgactactacttggaaacgtatcaaattggtacgtttaattcacgcc
Sequenced 30885    341  .....a...c.....

jgi-Phyci1-30885    401  actggaaccggagattctacaagaagtacgaggagtggtacaagcgaacg
Sequenced 30885    391  .....

jgi-Phyci1-30885    451  caccggactgggtttcggagataacgaaatga
Sequenced 30885    441  -----

```

Alignment of jgi-Phyci1-16230 ORF from the *Phytophthora cinnamomi* genome with the sequenced nucleotide data

```

jgi-Phyci1-16230      1  atgaagctttcgtacctcatcgcctttgccgccgtggctcgtcgcctcgac
Sequenced 16230      -----

jgi-Phyci1-16230      51  tgccgtccccgccagcgcctccacgggttgacgactaacaacctcgccg
Sequenced 16230      1  .....

jgi-Phyci1-16230     101  aggacttccaggtcgccccggagcttgctgccttcgcagcctccgcggc
Sequenced 16230      51  .....

jgi-Phyci1-16230     151  gccaacaggacagctccaagggcgatgacaagaagccaagggcgagca
Sequenced 16230     101  .....

jgi-Phyci1-16230     201  cggcgacaagaagaagggcggcgacaagaagaaggacggcgacaagaaga
Sequenced 16230     151  .....

jgi-Phyci1-16230     251  aggacggagacaagaagaagaaggaa-----aagaaggagaagaag
Sequenced 16230     201  .....raagaaggag.....

jgi-Phyci1-16230     292  gagaagaggaaggaggaagagagagagaagaagcacggcaagaaggacgg
Sequenced 16230     251  .....

jgi-Phyci1-16230     342  tgacaagaagcacggcaagaaggacggtgacaagaagcacggcaagaagg
Sequenced 16230     301  k...r...s.....k...r...s.....

jgi-Phyci1-16230     392  atggcgaaaagaagcacggcgacaagaagcacgagcagaaggacggcaag
Sequenced 16230     351  .....r.....r.....

jgi-Phyci1-16230     442  aaggaccacaagaagactgacggcagcaagaagtag
Sequenced 16230     401  .....yr.....

```


C

* 20 * 40 pfam05630: NPP1 * 60 *

Pc501748 : : 1
 U15350 : : -
 Pi_NPP1-like : MNLRALLFVVASLATVTADVVMIDHDAVQFFAQPEPTMESQKSAVKYKPKQLHISYGCHPYPAVQEDGSI : 70

80 * 100 * 120 * 140

Pc501748 : : 10
 U15350 : : 8
 Pi_NPP1-like : SAGLKWAGWPNGDCKGSALGSQVYSRSDWYKKGWAIMYAWYFPKSRQYVVDKYRFGHRHYWSYALVWITDNP : 140
 y166wi1NP

* 160 * 180 * 200 *

Pc501748 : : 79
 U15350 : : 77
 Pi_NPP1-like : EAENPAILGVVVRTAGGNERRTPPDAKF-IDGSSLKLDYYKSSWHGKTGLQLTENTGEFQDLITWEQMT : 210
 eaeNpaILGVw6r3agGne4R3PPdaK5 6DG336KLD5YkSsWhgKtGLQLTEh3GEfQDLITWEQ63E

220 * 240 *

Pc501748 : : 119
 U15350 : : 109
 Pi_NPP1-like : EARNALSEADFDVSWAFTRAV-MPLKDDIFTSRLKKAYPF : 249
 EARKALSEADFDghW erKd d6PfiDseFTg lkka f

D

* 20 * 40 * 60 *

Pc227507 : : 70
 U2360 : : 70
 Ps_cellulase : DLKDSLSTDATWGDGSDTDWSVAATELGNHLLTKCPKWLAFVQGVQGESHKDITYGSRSIKNTFLPGSDLT : 70
 DLKDsLSTDATWGDGSDTDWSVAAT2LGNHLLTKCPKWLAFVQGVQGESHKDITYGSRSIKNTFLPGSDL3

80 * 100

Pc227507 : : 100
 U2360 : : 100
 Ps_cellulase : GVSDAPITLDTDNKVVYSPKYSSSYLPR : 100
 GVSttPITLskDNKVVYSPKYSSSY6PRL

E

* 20 * 40 * 60 *

Pc324897 : : 54
 U3104 : : 75
 Pi_mucin : SSGDAAVGIAAVTDDSLYGLGCFNDHCRFCKVQNTTQSAEFVDCSSLVGFSSPSAAPVDVVTTL : 75
 q tcgi aa GDiaVG6hi TD 3Cs GG6GCiNDlCRFCK6 tT QS a56DC sl g 3s 3aAPVd tt

Pc324897 : - : -
 U3104 : P : 76
 Pi_mucin : P : 76
 p

F

pfam00964: elicitin

```

Pc96866      : -----MPVNVA FVVVSLALMGATNANDE CSTS QLMGLAGNKNVAGCSKAA GFSGM : 50
U4987       : AQFPTTGIFKMPVNVA FVVVSLALMGATNANDE CSTS QLMGLAGNKNVAGCSKAA GFSGM : 60
Pi_elicitin_INL3a-like : M-----MPSLTS FVLI GLALAGSVSAED -CSTEALMKLASSTNLAGCTADTGVS-V : 49
                MPvnvaFV66sLALmGatnAnDeCSTsqLMgLAgnkN6AGC3kaaGfSg6

Pc96866      : STIADLSPBQLKAVCGSSACVALMKDMAAMD LGNCRIFESKIH LQTDITDAFNEKCSARG : 110
U4987       : STIADLSPBQLKAVCG----- : 76
Pi_elicitin_INL3a-like : STIADLSPBQLKAVCGSSACVALMKDMAAMD LGNCRIFESKIH LQTDITDAFNEKCSARG : 109
                STIadL3peQ6kAVCGss c lm d a lg c ip s q di d cs g

Pc96866      : SMDG--SSSSTGSSVSSSNIRES SAASSASS-----ATTVVIGYAWAITL : 152
U4987       : ----- : -
Pi_elicitin_INL3a-like : TMGSM AAGSSSSSNGGTNVGDESGTSSSSSKATGSSSSTSGAGNTAGTIVTIGFASAV-- : 167
                m s ss s n s ss ss a tv ig a a

Pc96866      : AI VAMLL- : 159
U4987       : ----- : -
Pi_elicitin_INL3a-like : VA VAMILM : 175
                vam l
    
```

G

cd01100: APPLE factor VI-like

```

Pc89138 : MRVLTVAATAIAVFAVLSDAGCPNTSFGKCGDVNSPECCTDGNYCMPWTSSYQCLELPSQCARQFTGYDFYGGD : 75
U6179   : -----FTGYDFYGGD : 10
Pi_CBEL : MRVLTVAATAIAVVVGKSDAACPNINLGRCGDASNPECCPEGNYCMPWAANNYQCLELPSQCARQFTGYDFYGGD : 75
                mr ltvaataiav sda cpnt g cgd pecc gnycmpw yqclp p qc rqFTGYDFYGGD

Pc89138 : IKTVYGLQPGDCCATCLSTSGCLAYTFINEYQGSTACFLKAGMGQPRKVA GAISAVVDGYTSDQDHTPKRRLOGN : 150
U6179   : IKTVYGLQPGDCCATCLSTSGCLAYTFINEYQGSTACFLKAGMGQPRKVA GAISAVVDGYTSDQDHTPKRRLOGN : 85
Pi_CBEL : IKTVYGLQPGDCCATCLSTSGCLAYTFINEYQGTTACFLKAGMGQPRKVV GAMSAVLDSYTSQDQDHTPKRRLOGN : 150
                IKTVYGLQPGdCCATCLsTSGCLAYTFfINEyQG3TACFLKAGMGQPRKVaGA6SAV6DgYTSQDQDHTPKRRLOQ1

Pc89138 : EPRFEIPELS : 160
U6179   : E----- : 86
Pi_CBEL : SPRVELPGL- : 159
                epr e p l
    
```

H

pfam01083: cutinase

```

Pc6902      : -----*      20      *      40      *      60      *      : 50
U6180      : PINAPHRRTRPPHTIPVTTSTMKLTSAACSTILGSLLVLVANPTYLVNAECSDVHVVFARGSGEMAGLGICG : 71
Pi_cutinase : -----*      20      *      40      *      60      *      : 48
              MKLTSACSTILGSLLVLVANPTYLVNAECSDVHVVFARGSGEMAGLGICG
              MKLTSaC3tI6GSLlvL6anpty6VnA2CSDVHVVFARGSGEMAGLGICG

Pc6902      : -----*      80      *      100     *      120     *      140     : 121
U6180      : EPLVSGITSNLGGMSVSSYAVSYLASMDQTSAGPGATDMTNHVVSVAQECPNVTFVVLGGYSQGASVTDIAI : 142
Pi_cutinase : EPLVSGITSNLGGMSVSSYAVSYLASMDQTSAGPGATDMTNHVVSIAQECPNVTFVVLGGYSQGASVTDI : 119
              EPLVSGITSNLGGMSVSSYAVSYLASMDQTSAGPGATDMTNHVVS6AQECPNVTFVVLGGYSQGASVTDIa6

Pc6902      : -----*      160     *      180     *      200     *      : 192
U6180      : GIKTMLGSGDTIPEdLAPRIKAVVTFGNPLKLMGQTISSSSSTYGSKAIEFCNtGDPVCGNGFNvMAHMTY : 213
Pi_cutinase : GIKTMLGSGDTIPEdLAPRIKAVVTFGNPLKLMGQTISSSSSTYGSKAIEFCNtGDPVCGNGFNvMAHMTY : 190
              GIKTMLG3Gd3IPEdLAPRIKAVVTFGNPLKLMGQTISSSSSTYGSKAIEFCNtGDPVCGNGFNvMAHMTY

Pc6902      : -----*      220     *      240     *      260     : 217
U6180      : PMDSVSSAAEKAAALVKGGRTLRG----- : 261
Pi_cutinase : PMDSVPGAAEKAAALVKGSARMLRA----- : 215
              PMDSVssAAEKAAALVKGgRtLRg
  
```

I

pfam00295: glycosyl hydrolases family 28 includes polygalacturonase

```

Pc172110   : -----*      20      *      40      *      60      : -
U6659      : ----- : -
Pc_polygalacturonase : MKFFTALAVFALIATTANGSPMLRMEAEgKkGKSKTEAPVVEADDDYtQQDPTQQDPTQ : 62

Pc172110   : -----*      80      *      100     *      120     : 21
U6659      : ---NGADDEERAW-----GLKTL LSLVKT : 9
Pc_polygalacturonase : QDQSQQQDPTQQQSETTKTSSGCHLTGTYYKKGTSIASCSSIVIDSLTVPAGVTL LSLAKT : 124
              dp w p g lsl kt

Pc172110   : -----*      140     *      160     *      180     : 58
U6659      : -----*      140     *      160     *      180     : 59
Pc_polygalacturonase : GATIEfVGTTFfGTQKMEGPLVRVSGTSLT--vKGSgVLdGQGSWYwKQGSITRPVFFKlQ : 184
              f g w g v g v gsgvL gq sWyw g2sI r vffkl

Pc172110   : -----*      200     *      220     *      240     : 95
U6659      : ---ATLDLSKRKNIPYGSILQEFLNTRYLRQGAANLRRGN----- : 121
Pc_polygalacturonase : NVISSTVSGFTIKKMSFRtSSIVTSKdMMLRGLTIDNRtGNgiSKNTDSFDLTkNAHVTITG : 246
              nvis T6sgf iKn6p5r3 sivtskd LRgltilnR GNgi kntd fdltkn h titg

Pc172110   : -----*      260     *      280     *      300     *      : -
U6659      : NKIYNQDDYLAMQSSANTVFSN SYCCVg----- : 149
Pc_polygalacturonase : NKIYNQDDCLAMQSSANTVFSN SYCCGCHGISIGSLGGNAVDQSTTVQGLTVQGNtIVDSdN : 308
              nkiynqdd lamqss ntvfsn ycc g
  
```

J

```

                *           20           *           40           *           60           *
Pc98652 : SVLDSFDRRASLSARVTIEGPPEPFMRSAVSRKKLNSTVPRYLDYENAPGYAERAQKQYERRKRLEEFNAAK : 72
U6809   : SVLDSFDRRASLSARVTIEGPPEPFMRSAVSRKKLNSTVPRYLDYENAPGYAERAQKQYERRKRLEEFNAAK : 72
Pi_adhesin : SVLDSVDRRASLSARAAIDGPPEPFIRSAVSRKKLNSTLPRYLDYENSSGYAERAARKQYERRKRLEEFNAAK : 72
          SVLDSfDRRASLSARvtIeGPPEPF6RSVSRKKLNST6PRYLDYENapGYAERAqKQYERRKRLEEFNAAK

                80           *           100           *           120           *           140
Pc98652 : SERRQKELRVFFSERQOKALKLSAEVRRGLEAHEFAKLAKESEIEVQKKLRKEKQRELTARTHSRGSSTTS : 144
U6809   : SERRQKELRVFFSERQOKALKLSAEVRRGLEAHEFAKLAKESEIEVQKKLRKEKQRELTARTHSRGSSTTS : 144
Pi_adhesin : SEQRQRELRTFFNDRQOKSMLSAAEVRRGLEAHEFAQLAKESEIEAOKARRRGOQTRPTRASQRALSTTS : 144
          SErRQ4ELRvFFseRQQKa6klSAEVRRGLEAHEFAkLAKESEiEvQKk1R4ekQreltaRthsRgsS3TS

                *           160           *           180
Pc98652 : SVGASSNGLASRSSKKS--SASSVHKKIMTAPVAELTELETDPA : 187
U6809   : SVGASSNGLASRSSKKS--SASSVHKKIMTAPVAELTELETDPA : 187
Pi_adhesin : ---ASSSPGVASRTSNKSTVSSASVVEKVVAAVETASTVVTTE-- : 184
          svgASSSnG6ASR3SkKS SasSVhkK66tApVael3e6eTdpa

```

K

```

                *           20           *           40           *           60
Pc96960 : ----MKLSVFFV--AALLGVAALPAADASLHLRVHILTEKAGLYQQCEWEDKAVKCDTGMFCQMKEKHF : 63
U8663   : SSVAMKLSVFFV--AALLGVAALPAADASLHLRVHILTEKAGLYQQCEWEDKAVKCDTGMFCQMKEKHF : 67
Pi_SCR108-like : ----MKLVLI--AALLSLAAPVADASLHLRVHILTEKAGLYQQCEWENKAVKCEAGMFCQMKEKHF : 64
          MKLsVf6 AALLg6AALPaADASLHLRVHILTEKAGLYQQCEWE1KAVKCDtGMFCQMKEKHF

                *           80           *           100           *           120           *
Pc96960 : WCQKQKPLNDQCQGGKGDGPWAVPCSGGTNLKCVLISDQYSKCQKTDREKIKMPSNHHKGA--CAQQ : 131
U8663   : WCQKQKPLNDQCQGGKGDGPWAVPCSGGTNLKCVLISDQYSKCQKTDREKIKMPSNHHKGA----- : 130
Pi_SCR108-like : WCMKKSPLNDQCQGGKSTDGPWAVPCS--DNLKCVLISDQYSKCQKTDREKVKMP--NHHKGA--GVEE : 129
          WCqKqkPLNDQCQGGKgvDGPWAVPCsgg3NLKCVLISDQYSKCQqKTDREK6KMPsNHHKGA g

                140           *           160           *           180           *           200
Pc96960 : EKVPLNGRCKFEDGRKSCASGLQCIEDSEWSGNCKMKKEAGLYDQCAGQELRGLWKA--CPKGALCRES : 199
U8663   : ----- : -
Pi_SCR108-like : EKVALNERCKFENCTKGCASGLQCVEDSDWTGCMCKKEANVYDQCAGQEIIRGLWKAMCPKGAL---- : 192
          ekv ln rckfe g k casglqc eds w g cmkkea ydqcagqe irglwka cpkgal

                *           220
Pc96960 : TFYSQCLPEAVSLAMDA : 216
U8663   : ----- : -
Pi_SCR108-like : ----- : -

```

L

```

                *      20      *      40      *      60      *
Pc103247 : -----TRRPR : 5
U8949 : TTGSGGSLETTTTGSGATKTTVTTGSGASTTTTGTRTSTSTNSGTSSTQTTTSGSSNKGSSSTSTSTSTGKTPG : 74
Pi_mucin : TTGSGGSLETTTTGSGATTTIVTTGSGESTTTTSGTSSSTGSHSTHHTT-----TESSASGSTGTETPG : 63
          ttgsggsliettttgsgat ttvttgsg sttt gt st s t tt s s stgT tPg

                80      *      100      *      120      *      140
Pc103247 : R-----STPAATKPSATPAATTPSATPAATTkQSETCIDVSVEG : 44
U8949 : ATPAATTPGATPAATTPGATPAATKPS-----STPAATKPSATPAATTPSATPAATTkQSETCIDVSVEG : 139
Pi_mucin : ATPAATTPGATPAATTPGATPAASKPSATPAATTAGATPAASKPSATPAATTPSATPAATTEQSKTCIDVSVEG : 137
          atpaattpgatpaattpgatpaa kps sTPAA3KPSATPAATTPSATPAATTkQSeTCiDVSVEG

                *      160      *      180      *      200      *      220
Pc103247 : DATYCIIEGPICSGSGDLPAgkLCPVKGdVAVQDCHENLSSYTNSSTCVAPQDAVCMKIKTGAWG----- : 109
U8949 : DATYCIIEGPICSGSGDLPAgkLCPVKGdVAVQDCHENLSSYTNSSTCVAPQDAVCMKIKTGAWGCVFEK-NATPA : 213
Pi_mucin : DATYCIIEGPICSGAGDLPAgSLCPKKGdEAVQDCHDTLSSYTNSSTCVAPQDAVCLKIKTGAWGCVFEK-ATPA : 210
          DATYCIIEGPICSGsGDLPAgkLCPvKGdVAVQDCHenLSSYTNSSTCVAPQDAVC6KIKTGAWGcvfek atpa

                *      240      *      260      *      280      *
Pc103247 : ----- : -
U8949 : PTTTETGKATPAPTK-TEKSTPAPTTTETEKPPTPAPTAEIKINAPTPAPT--TTTGGKATPAP-----TTKES : 279
Pi_mucin : PTTTATDESTPAPTKTTDESTPAPTTAETGAPTPAPTKETENKAPTPAPTKQETEMDKPTPAETKPKETTTDKP : 284
          pttt t tpaptk t stpaptt et ptpapt et aptpapt t k tpa p tt

                300      *      320      *      340      *      360      *
Pc103247 : -----A : 110
U8949 : TPAPT--TTTGGKATPAPTtKESTPAPTTTEEATPAPT-----TKESTPAPTtKESTPAPTtPCEKTPEATKT : 346
Pi_mucin : TPAPTtKPEtETGKPTPAPTtKEGStPAPTTAGPStEAPTtEVPStESPtPKPTtETStDAPtPStGVPSIDG : 357
          tpapt t tgg tpapt stpaptt t apt t tp ptt st ap tp p t

                380      *      400
Pc103247 : PTPCPTtEVPStEAPTtPCPStE---VPStEAPTtCPT : 144
U8949 : StEAPTtAGPStEAPTtCPTtE---VPStGAPTtCPT : 380
Pi_mucin : StPStPttEVPtTDAPTtPCPttEVPtVPTTDAPTtPPT : 394
          sTP PTTevP3TeAPTtPCP3TE VP3T APTtCPT

```

Online Resource 4a. Transcript specific primers were designed from the assembled transcriptome and used to amplify and sequence full length transcripts to assess the robustness of the assembly. Transcript sequences were aligned with the nucleotide sequences from the *Phytophthora cinnamomi* genome.

Online Resource 4b. Protein sequence alignment of transcripts from the cyst and germinating cyst cDNA library of *Phytophthora cinnamomi*, to the best blastx hit in the genome of *P. cinnamomi* (top sequence) and the *Phytophthora* protein sequence used to annotate the selected transcript. Conserved domain names are shown above the relevant domain sequence (blocks). Shaded amino acids indicate amino acids with 100% conservation (black) or only 80% conservation between two or three of the sequences (grey) (GeneDoc Version 2.7.000, conservation mode with similarity groups enabled). The consensus sequence is indicated underneath each alignment. Note only the complete transcript sequence is shown in the figure but not the whole protein sequence of the protein it is aligned to. A) C31843 aligned to Phyci1-16230 and Physo_298800 B) C83210 aligned to Phyci1-246465 and Phyra_97565 C) U15350 aligned to Phyci1-501748 and PITG_19936 D) U2360 aligned to Phyci1-227507 and EGZ23444.1 E) U3104 aligned to Phyci1-324897 and PITG_15968 F) U4987 aligned to Phyci1-96866 and PITG_06370 G) U6179 aligned to Phyci1-89138 and PITG_03640 H) U6180 aligned to Phyci1-6902 and PITG_12361 I) U6659 aligned to Phyci1-98110 and AAN05466.1 J) U6809 aligned to Phyci1-98652 and PITG_16183 K) U8863 aligned to Phyci1-96960 and PITG_16145 L) U8949 aligned to Phyci1-103247 and PITG_09760.

*Note: the annotations of these transcripts were confirmed through blastx analysis (E-value $>1 \times 10^5$) to *Phytophthora* proteins in the National Centre for Biotechnology Information (NCBI) database (<http://www.ncbi.nlm.nih.gov/>), FungiDB database version 2.2 (Stajich et al. 2012), the *Phytophthora infestans* database (*Phytophthora infestans* Sequencing Project, Broad Institute of Harvard and MIT (<http://www.broadinstitute.org/>), downloaded February,

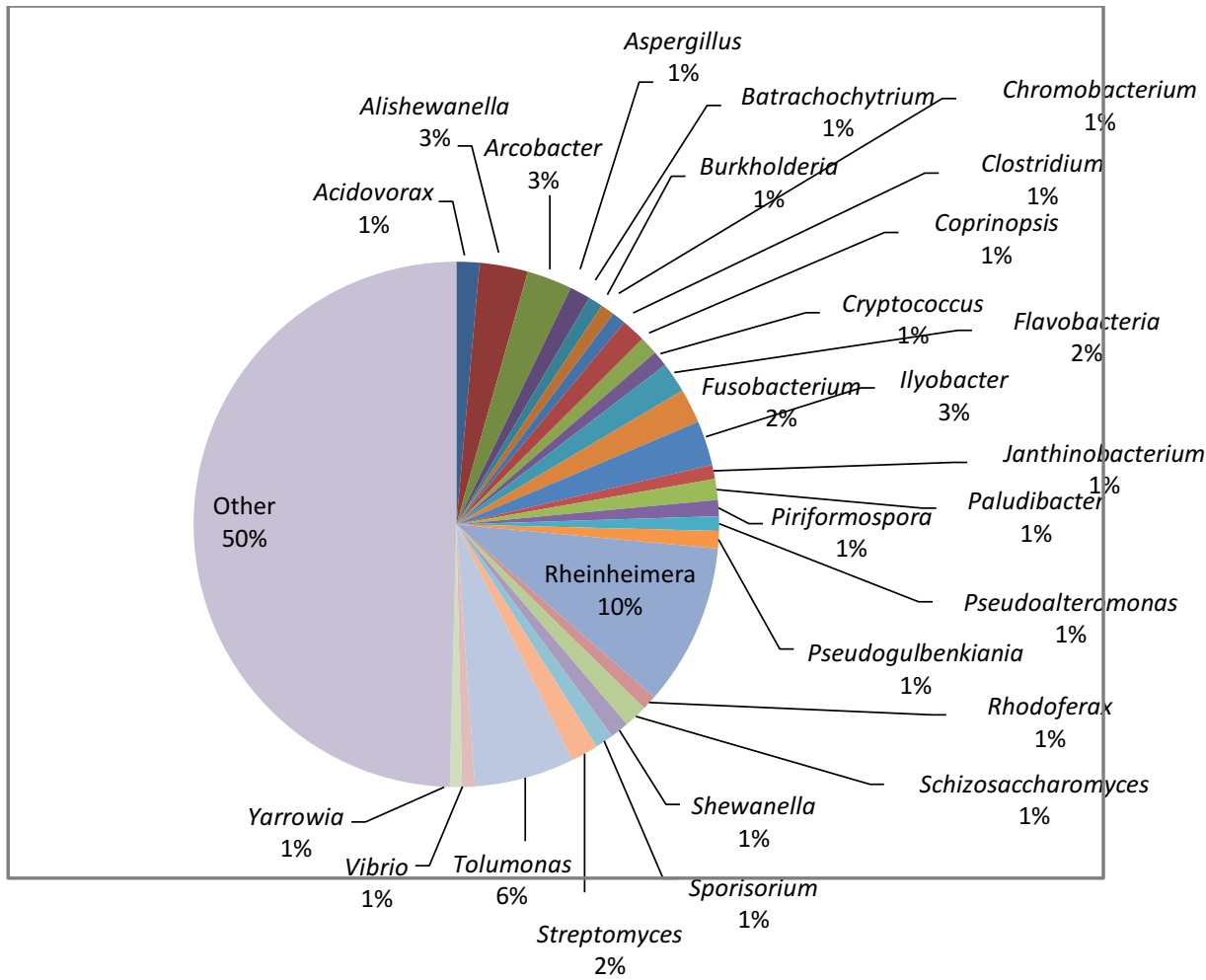
2013) and the *P. cinnamomi* genome v1.0

(<http://genome.jgi.doe.gov/Phyci1/Phyci1.home.html>, downloaded March, 2013).

Online Resource 5. Transcripts from the *Phytophthora cinnamomi* cyst and germinating cyst RNA-seq library (CNVDB29), with the best protein match from the *Phytophthora cinnamomi* genome indicated as determined through blastx analysis (E-value 1×10^5).

Unigene ID	Protein ID of best blast hit	E-value	Bit Score	% Identity	% Similarity	% Gaps	Alignment length (aa)
C31843	jgi-Phyci1-16230	1.00E-14	65.47	100	100	0	95
C83210*	jgi-Phyci1-246465	3.09E-35	133.26	100	100	0	82
U15350*	jgi-Phyci1-501748	8.28E-64	228.79	98	99	0	109
U2360	jgi-Phyci1-227507	1.14E-48	178.33	100	100	0	100
U3104	jgi-Phyci1-324897	1.52E-23	94.36	39	52	28	54
U4987*	jgi-Phyci1-96866	6.11E-35	132.11	86	86	13	66
U6179*	jgi-Phyci1-89138	2.57E-51	186.81	100	100	0	86
U6180	jgi-Phyci1-6902	8.44E-115	399.82	83	83	16	217
U6659*	jgi-Phyci1-98110	7.95E-56	192.2	6	11	57	91
U6809*	jgi-Phyci1-98652	6.63E-76	270.01	100	100	0	187
U8663	jgi-Phyci1-96960	4.28E-65	233.42	96	96	3	126
U8949*	jgi-Phyci1-103247	1.25E-45	170.63	35	35	62	144

* Unigenes for which the expression pattern was determined in this study.



Online Resource 6. Distribution of the genera represented by *Phytophthora cinnamomi* transcripts regarded as contaminant sequences, as annotated in the NCBI non redundant protein database. Transcript functional annotations are based on protein sequence similarity as determined with blastx analysis (- -value $>1 \times 10^5$) to the NCBI non-redundant protein database. Approximately 42 % of the transcripts showed significant similarity to proteins originating from the domain bacteria and 8.7 % from the domain fungi. Genera that presented less than 1% of the dataset were not included in the analysis and are designated as 'other genera' in the figure (50 %).

Online Resource 7. Putative secreted proteins identified from *Phytophthora cinnamomi* CNVDB29 cysts and germinating cysts with their best hit from the genomes of *Phytophthora infestans* and *Phytothora cinnamomi* CBS144.22 as determined through blastx analysis (E-value $<1 \times 10^{-5}$). Attached PDF file.

Online Resource 8. Complete list of transcripts with a putative role in *Phytophthora cinnamomi* pathogenicity according to the literature. Attached - xcel file.

Online Resource 9. Putative RXLR effector transcripts identified from *Phytophthora cinnamomi* CNVDB29 with the corresponding protein from the *P. cinnamomi* genome (CBS 144.22) as determined through blastp analysis (E-value $<1 \times 10^{-5}$).

Query ID ^a	SP	RXLR-motif	E-value	Score (bit)	Hit ID ^b
C3889	No	Yes	0	758	Phyci1-241851
C14564	No	No	3.72E-143	503	Phyci1-88880
C38826	No	Yes	1.42E-100	361	Phyci1-113905
C12983	No	Yes	1.08E-81	299	Phyci1-184564
C14063	Yes	Yes	1.41E-79	291	Phyci1-30885
C29027	No	Yes	5.19E-70	260	Phyci1-92480
C17530	Yes	Yes	3.01E-69	257	Phyci1-292291
C28374	No	Yes	1.11E-59	224	Phyci1-85009
C46416	No	Yes	5.14E-57	215	Phyci1-400789
C33615	Yes	Yes	5.57E-57	215	Phyci1-80513
C170656	No	Yes	1.67E-52	200	Phyci1-91700
C15073	Yes	Yes	1.07E-51	198	Phyci1-172110
C17531	Yes	Yes	1.37E-47	184	Phyci1-292291
C66803	Yes	Yes	1.35E-44	173	Phyci1-557761
C19961	No	Yes	5.34E-44	172	Phyci1-20391
C33954	No	Yes	5.10E-35	142	Phyci1-214852
C83210	Yes	Yes	3.59E-34	139	Phyci1-246465
C129865	Yes	No	4.44E-31	129	Phyci1-23225
C5551	Yes	Yes	1.77E-27	117	Phyci1-442085
C17839	No	Yes	3.32E-25	111	Phyci1-7143
C84507	Yes	Yes	1.24E-21	97	Phyci1-381040
C31843	Yes	Yes	1.30E-13	70	Phyci1-16230
C118237	No	No	2.61E-13	70	Phyci1-95101
C78872	No	Yes	1.93E-10	60	Phyci1-492221
C112596	No	Yes	6.70E-10	58	Phyci1-97937

C19412	No	Yes	4.20E-01	29	Phyci1-84057
C22794	Yes	No	9.90E-01	31	Phyci1-256563
C43424	No	Yes	1.73E+00	28	Phyci1-115216
C1990	Yes	Yes	2.78	29	Phyci1-228636
C8139	Yes	Yes	2.91	29	Phyci1-303442
C2637	No	Yes	3.74	26	Phyci1-69046
C2634	No	Yes	3.76	26	Phyci1-69046
C14204	Yes	Yes	7.49E+00	27	Phyci1-222893
C2006	No	Yes	8.23	27	Phyci1-65069
C9909	No	No	NA	NA	NA
C145768	No	Yes	NA	NA	NA
C146670	No	Yes	NA	NA	NA
C1480	No	Yes	NA	NA	NA
C17241	No	Yes	NA	NA	NA
C5149	No	Yes	NA	NA	NA
C2921	Yes	Yes	NA	NA	NA
C31644	Yes	Yes	NA	NA	NA
C33937	Yes	Yes	NA	NA	NA
C81356	Yes	Yes	NA	NA	NA

SP – signal peptide

^aIdentity numbers from *P. cinnamomi* CNVDB29 as explained under methods, functional annotation.

^bProtein identity numbers from the *P. cinnamomi* genome at <http://genome.jgi.doe.gov/>, downloaded March, 2013

References

- King, A. J., Reeve, W., Van der Hoek, G. B., Williams, N., McComb, J., O'Brien, P. A., et al. (2010). Defining the phosphite-regulated transcriptome of the plant pathogen *Phytophthora cinnamomi*. *Molecular Genetics and Genomics*, 284(6), 425-435, doi:10.1007/s00438-010-0579-7.
- Kong, P., Hong, C. X., & Richardson, P. A. (2003). Rapid detection of *Phytophthora cinnamomi* using PCR with primers derived from the *Lpv* putative storage protein genes. *Plant Pathology*, 52(6), 681.
- Narayan, R. D., Blackman, L. A., Shan, W., & Hardham, A. R. (2010). *Phytophthora nicotianae* transformants lacking dynein light chain 1 produce non-flagellate zoospores. *Fungal Genetics and Biology*, 47(8), 663-671.
- Stajich, J., Harris, T., Brunk, B., Brestelli, J., Fischer, S., Harb, O., et al. (2012). FungiDB: an integrated functional genomics database for fungi. *Nucleic Acids Research*, 40, D675-D681.